AD-A105 413   TEXAS A AND M UNIV COLLEGE STATION INST OF STATISTICS   F/G 12/1
THE FINITE MEMORY PREDICTION OF COVARIANCE STATIONARY TIME SERI--ETC(U)
APR 81   H J NEWTON, M PAGANO   DAAG29-80-C-0070
UNCLASSIFIED   TR-N-21   NL

1 OF 1
AD A
105413

END
DATE
FILMED
10-81
DTIC

TEXAS A&M UNIVERSITY

COLLEGE STATION, TEXAS 77843

INSTITUTE OF STATISTICS
Phone 713 - 845-3141

LEVEL

THE FINITE MEMORY PREDICTION OF COVARIANCE STATIONARY TIME SERIES

~~by~~ H. J. Newton ~~and~~ Marcello Pagano

Institute of Statistics — Texas A&M University

Statistical Science Division — State University of New York at Buffalo

Technical Report No. N-21

April 1981

OCT 1 9 1981

Texas A & M Research Foundation
Project No. 4226T

"Robust Statistical Data Analysis and Modeling"

Sponsored by the Office of Naval Research

Professor Emanuel Parzen, Principal Investigator

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER N-21	2. GOVT ACCESSION NO. AD-A105 413	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Finite Memory Prediction of Covariance Stationary Time Series		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) H.J. Newton and Marcello Pagano		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0070 ONR N0001481MP10001
9. PERFORMING ORGANIZATION NAME AND ADDRESS Texas A&M University Institute of Statistics College Station, TX 77843		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office P.O. Box 12277 Research Triangle Park, NC 27709		12. REPORT DATE April
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Covariance stationary time series: minimum mean square linear prediction; modified Cholesky decomposition algorithm; autoregressive-moving average time series.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An algorithm is presented for conveniently calculating h step ahead minimum mean square linear predictors and prediction variances given a finite number of observations from a covariance stationary time series Y. It is shown that elements of the modified Cholesky decomposition of the covariance matrix of observations play the role in finite memory prediction that the coefficients in the infinite order moving average representation of Y play in infinite memory prediction. The algorithm is applied to autoregressive-moving average time series where further simplifications are shown to occur. A numerical example illustrating the basic points of the general algorithm is presented.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

The Finite Memory Prediction of Covariance Stationary Time Series

by H.J. Newton and Marcello Pagano

Institute of Statistics, Texas A & M University

and

Harvard University and Sidney Farber Cancer Institute

Summary

An algorithm is presented for conveniently calculating h step ahead minimum mean square linear predictors and prediction variances given a finite number of observations from a covariance stationary time series Y. It is shown that elements of the modified Cholesky decomposition of the covariance matrix of observations play the role in finite memory prediction that the coefficients in the infinite order moving average representation of Y play in infinite memory prediction. The algorithm is applied to autoregressive-moving average time series where further simplifications are shown to occur. A numerical example illustrating the basic points of the general algorithm is presented.

Some key words: Covariance stationary time series: minimum mean square linear prediction; modified Cholesky decomposition algorithm; autoregressive-moving average time series.

1. Introduction

As pointed out by Whittle (1963, p.47), the calculations required to find finite memory predictors for covariance stationary time series are made difficult by the need to calculate the inverse of the $T \times T$ covariance matrix of the observations.

Thus many authors (see Box and Jenkins (1970, p. 126) have proposed using approximate infinite memory predictors rather than finding the exact finite memory predictors.

Pagano (1976) has given an algorithm for finite memory prediction of a pure moving average process which reduces much of the calculation in the general algorithm. Ali (1977) uses a well known result to reduce inverting the $T \times T$ matrix to the successive inversion of smaller matrices.

The purpose of this paper is to propose a general algorithm for prediction of covariance stationary time series which capitalizes on the special structure of the modified Cholesky decomposition of a symmetric Toeplitz covariance matrix. Section 2 contains the algorithm as theorem 1 which also shows the analogy of the algorithm with infinite memory prediction. In section 3 theorem 2 presents the results of applying theorem 1 to autoregressive-moving average processes. Finally a numerical example is presented in section 4 illustrating theorem 1.

2. Finite Memory, Horizon h, Minimum Mean Square Linear Prediction of Covariance Stationary Time Series.

Consider a zero mean covariance stationary time series $\{Y(t), t=0,\pm 1,\ldots\}$ with autocovariance function $R(v) = E(Y(t)Y(t+v))$. Then given observations $Y(1), \ldots, Y(T)$, the horizon h, memory T, minimum mean square error linear predictor $Y(T+h|T)$ of $Y(T+h)$ is given by that linear combination of $Y(1), \ldots, Y(T)$ that minimizes $E\{Y(T+h) - Y(T+h|T)\}^2$.

Thus

$$Y(T+h|T) = \sum_{j=1}^{T} \lambda_{T,h}(j)\, Y(T+1-j)$$

where $\underset{\sim}{\lambda}_{T,h} = (\lambda_{T,h}(1), \ldots, \lambda_{T,h}(T))^T$ satisfies

$$\Gamma_T \, \boldsymbol{\lambda}_{T,h} = \boldsymbol{r}_{T,h} \;, \quad \text{where } \boldsymbol{r}_{T,h} = (R(h), \ldots, R(h+T-1))^T$$

and $\Gamma_T = \text{TOEPL}\,(R(0), \ldots, R(T-1))$, <u>i.e.</u> Γ_T is the $T \times T$ symmetric Toeplitz matrix having (j,k) element $R(|j-k|)$.

Suppose that Y is purely nondeterministic, <u>i.e.</u>

$\sigma_\infty^2 = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\omega)d\omega\right\} > 0$ where f is the spectral density function of Y. Then $Y(t)$ can be represented as the limit in mean square of an infinite order moving average process, <u>i.e.</u>

$$Y(t) = \sum_{k=0}^{\infty} \beta_\infty(k)\, \varepsilon(t-k) \tag{1}$$

where $\varepsilon(t)$ is the infinite memory horizon one error in predicting $Y(t)$ and $E(\varepsilon(T)\varepsilon(T-j)) = \delta_j \, \sigma_\infty^2$ for all integer T and j, where δ_j is the Kronecker delta. Also the horizon h, minimum mean square error linear infinite memory predictor $Y(T+h|T, T-1, \ldots)$ and prediction variance $\sigma^2_{T,h,\infty}$ are given by

$$Y(T+h|T, T-1, \ldots) = \sum_{k=h}^{\infty} \beta_\infty(k)\varepsilon(T+h-k) \;, \tag{2}$$

$$\sigma^2_{T,h,\infty} = \sigma_\infty^2 \sum_{k=0}^{h-1} \beta_\infty^2(k) \;. \tag{3}$$

The process Y being purely nondeterministic also means that its autocovariance function is positive definite. Thus for all T we can form the modified Cholesky decomposition (Wilkinson (1965)) $\Gamma_T = L_T D_T L_T^T$ of Γ_T where L_T is a $T \times T$ unit lower triangular matrix and D_T is a $T \times T$ diagonal matrix. An important property of L_T and D_T is that they are nested for increasing T, <u>i.e.</u>

$$L_{T+1} = \begin{bmatrix} L_T & 0 \\ \underset{\sim}{\ell}^T & 1 \end{bmatrix} , \quad D_{T+1} = \begin{bmatrix} D_T & 0 \\ 0 & d_{T+1} \end{bmatrix} .$$

Thus the (j,k)th element of L_T will be referred to as L_{jk}.

The following theorem shows the role played by L_T and D_T in finite memory prediction.

Theorem 1

Let Y be a purely nondeterministic covariance stationary time series with covariance function R. Let $\Gamma_T = L_T D_T L_T^T$ be the modified Cholesky decomposition of the covariance matrix of $\underset{\sim}{Y}_T^T = (Y(1), \ldots, Y(T))$. Define $\underset{\sim}{e}_T^T = (e(1), \ldots, e(T))$ by $L_T \underset{\sim}{e}_T = \underset{\sim}{Y}_T$. Then

a) $$Y(T+h|T) = \sum_{k=h}^{T+h-1} L_{T+h,T+h-k}\, e(T+h-k)$$

b) $$\sigma^2_{T,h} = E\{Y(T+h) - Y(T+h|T)\}^2 = \sum_{k=0}^{h-1} L^2_{T+h,\, T+h-k}\, d_{T+h-k}$$

c) i) $$\lim_{T\to\infty} L_{T,T-j} = \beta_\infty(j)$$

ii) $$\lim_{T\to\infty} d_T = \sigma^2_\infty$$

Proof of Theorem

a) Defining the T × T permutation matrix P_T to be a matrix of zeros with ones on the main reverse diagonal, we have $Y(T+h|T) = \underset{\sim}{\lambda}^T_{T,h} P_T \underset{\sim}{Y}_T$ where $\Gamma_T \underset{\sim}{\lambda}_{T,h} = \underset{\sim}{r}_{T,h}$ since premultiplication (postmultiplication) by P_T reverses row (column) order of a matrix. Thus $Y(T+h|T) = \underset{\sim}{r}^T_{T,h} \Gamma_T^{-1} P_T \underset{\sim}{Y}_T$ $= \underset{\sim}{r}^T_{T,h} P_T \Gamma_T^{-1} P_T P_T \underset{\sim}{Y}_T = \underset{\sim}{\rho}^T_{T,h} \Gamma_T^{-1} \underset{\sim}{Y}_T$ where $\underset{\sim}{\rho}_{T,h} = P_T \underset{\sim}{r}_{T,h}$ since $P_T^2 = I_T$ and since for the symmetric Toeplitz matrix Γ_T^{-1} we have $P_T \Gamma_T^{-1} P_T = \Gamma_T^{-1}$.

Thus $Y(T+h|T) = \underset{\sim}{\rho}^T_{T,h}\ L_T^{-T}\ D_T^{-1}\ L_T^{-1}\ \underset{\sim}{Y}_T = \underset{\sim}{\rho}^T_{T,h}\ L_T^{-T}\ D_T^{-1}\ \underset{\sim}{e}_T$. To show that this is the result in (a) we note 1) $\underset{\sim}{\rho}^T_{T,h} = (R(T+h-1), \ldots, R(h))$ is the last row of $\underset{\sim}{\Gamma}_{T+h}$ without its last h elements, 2) $\Gamma_T L_T^{-T}\ D_T^{-1} = L_T$ for all T, and 3) because of the nesting of the L_T and D_T, $L_T^{-T}\ D_T^{-1}$ is the $T \times T$ principal minor of the upper triangular matrix $L_{T+h}^{-T}\ D_{T+h}^{-1} = \Gamma_{T+h}^{-1}\ L_{T+h}$.

Thus $(\underset{\sim}{\rho}^T_{T,h}\ L_T^{-T}\ D_T^{-1})_k = (\Gamma_{T+h}\ L_{T+h}^{-T}\ D_{T+h}^{-1})_{T+h,k} = (\Gamma_{T+h}\ \Gamma_{T+h}^{-1}\ L_{T+h})_{T+h,k}$ $= L_{T+h,k}$, proving (a).

To prove (b), note that $\sigma^2_{T,h} = R(0) - \underset{\sim}{r}^T_{T,h}\ \Gamma_T^{-1}\ \underset{\sim}{r}_{T,h} = R(0)$ $-\underset{\sim}{\rho}^T_{T,h}\ L_T^{-T}\ D_T^{-1}\ L_T^{-1}\ \underset{\sim}{\rho}_{T,h} = R(0) - \underset{\sim}{\ell}^T_{T,h}\ D_T\ \underset{\sim}{\ell}_{T,h}$ where $\underset{\sim}{\ell}^T_{T,h} = \underset{\sim}{\rho}^T_{T,h}\ L_T^{-T}\ D_T^{-1}$ which as above is the row vector $(L_{T+h,1}, \ldots, L_{T+h,T})$. Also $R(0)$ $= \Gamma_{T+h,\ T+h} = (L_{T+h}\ D_{T+h}\ L^T_{T+h})_{T+h,\ T+h} = \sum_{k=1}^{T+h} L^2_{T+h,k}\ d_k$, thus proving (b).

To prove (c) we first note that multiplying both sides of (1) for $t = T$ by $\varepsilon(T-j)$ and taking expectations gives

$$E(Y(T)\varepsilon(T-j)) = \beta_\infty(j)\sigma_\infty^2$$

We next note that

$$E(e(T)e(T-j)) = \delta_j\ d_T\ ,$$

$$E(Y(T)e(T-j)) = \sum_{k=1}^{T} L_{T,k}\ E(e(k)e(T-j)) = L_{T,T-j}\ d_{T-j}$$

and that $e(1) = Y(1)$, $e(t) = Y(t) - Y(t|t-1, \ldots, 1)$, $t = 2, \ldots, T$, where the notation $Y(t|t-1, \ldots, 1)$ makes explicit which Y's are used in predicting $Y(t)$. Then by the stationarity of Y we have that $E(Y(T)e(T-j))$ $= E(Y(T)\ [Y(T-j) - Y(T-j|T-j-1, \ldots, 1)]) = E(Y(0)[Y(-j)-Y(-j|-j-1,\ldots,1-T)])$ which by a standard martingale convergence argument converges to

$E(Y(0)\varepsilon(-j)) = \beta_\infty(j)\sigma_\infty^2$. A similar argument shows $d_T \to \sigma_\infty^2$ thus proving (c).

Thus comparing (a) with (2) and (b) with (3), it is clear that the elements of L_T and D_T are playing the role in finite memory prediction of σ_∞^2 and $\beta_\infty(\cdot)$ in the infinite memor" case, while (c) makes explicit the connection. In the next section we describe how this algorithm can be simplified for an autoregressive-moving average process.

3. Application to Autoregressive-moving Average Processes.

The univariate autoregressive-moving average process $\{Y(t),\ t=0,\pm 1,\ \ldots\}$ of order (p,q) is defined by

$$\sum_{j=0}^{p} \alpha(j)Y(t-j) = \sum_{k=0}^{q} \beta(k)\varepsilon(t-k) \ , \quad t = 0, \pm 1, \ldots$$

where $\alpha(0) = \beta(0) = 1$, and $E(\varepsilon(t)) = 0$, $E(\varepsilon(t)\varepsilon(t+v)) = \delta_v\sigma^2$.

We assume that the zeros of the complex polynomial $g(z) = \sum_{j=0}^{p} \alpha(j)z^j$ are all greater than one in modulus so that Y does indeed have an infinite order moving average representation and that (defining $R_Y(v) = E(Y(t)Y(t+v))$

$$\sum_{j=0}^{p} \alpha(j)\ R_Y(j-v) = 0 \ , \quad v > q \ .$$

Then given a realization $\underset{\sim}{Y}_T = (Y(1),\ \ldots,\ Y(T))^T$ from $Y(\cdot)$ we define the following quantities:

i) $\Gamma_{Z,T}$ = TOEPL $(R_Z(0),\ \ldots,\ R_Z(T-1)$ where $Z(\cdot)$ is an autoregressive process of order p with coefficients $\alpha(1),\ \ldots,\ \alpha(p)$. Thus $Z(\cdot)$ is referred to as the autoregressive part of Y.

ii) $\underset{\sim}{X}_T = (X(1),\ \ldots,\ X(T))^T = V_{Z,T}\ \underset{\sim}{Y}_T$ where $\Gamma_{Z,T} = L_{Z,T}D_{Z,T}L_{Z,T}^T$ and $V_{Z,T} = L_{Z,T}^{-1}$.

Then it is well known that the j^{th} row of $V_{Z,T}$ is given by

$$\underset{\sim}{v}_j^T = \begin{cases} (1, \underset{\sim}{0}_{T-1}^T) & j = 1 \\ (\alpha_{j-1}(j-1), \ldots, \alpha_{j-1}(1), 1, \underset{\sim}{0}_{T-j}^T) , & j = 2, \ldots, p \\ (\underset{\sim}{0}_{j-p-1}^T \alpha(p), \ldots, \alpha(1), 1, \underset{\sim}{0}_{T-j}^T) , & j = p+1, \ldots, T \end{cases}$$

where $\sum_{\ell=0}^{k} \alpha_k(\ell) R_Z(\ell-v) = 0, \quad v = 1, \ldots, k < p$.

Thus there are only $p(p+1)/2$ distinct elements of $V_{Z,T}$ (other than 0 and 1) and $\alpha_j(k)$, $1 \leq k \leq j < p$ are easily obtained from $\alpha(1), \ldots, \alpha(p)$ by performing Durbin's recursive algorithm (1960) for decreasing j.

iii) $\Gamma_{X,T} = E(\underset{\sim}{X}_T \underset{\sim}{X}_T^T) = V_{Z,T} \Gamma_{Y,T} V_{Z,T}^T$ where $\Gamma_{Y,T} = \text{TOEPL}(R_Y(0),\ldots,R_Y(T-1))$.

Since for $t > p$, $X(t) = \sum_{j=0}^{p} \alpha(j) Y(t-j) = \sum_{k=0}^{q} \beta(k)\varepsilon(t-k)$, we have for

$j,k > p$, $\quad (\Gamma_{X,T})_{jk} = R_X(|j-k|)$

where

$$R_X(v) = \begin{cases} \sigma^2 \sum_{k=0}^{q-|v|} \beta(k)\beta(k+|v|) & , \ |v| \leq q \\ 0 & , \ |v| > q \end{cases} .$$

Thus $\Gamma_{X,T}$ is symmetric band Toeplitz in its last T-p rows and columns while its $p \times p$ principal minor is given by $V_{Z,p} \Gamma_{Y,p} V_{Z,p}^T$. Thus $\Gamma_{X,T}$ is almost the $T \times T$ covariance matrix of a pure moving average process.

iv) $\Gamma_{X,T} = L_{X,T} D_{X,T} L_{X,T}^T$. Since $(\Gamma_{X,T})_{j,k} = 0$ for $|j-k| > q$, then $L_{X,j,k} = 0$ also for $j-k > q$. Since $L_{X,T}$ is nested for increasing

T we refer to the (j,k)th element of $L_{X,T}$ for any $T \geq j,k$ as $L_{X,j,k}$.

v) The vector $\underset{\sim}{e}_T = (e(1), \ldots, e(T))^T$ by $L_{X,T}\underset{\sim}{e}_T = \underset{\sim}{X}_T$, i.e.

$$e(1) = X(1) \text{ , while } e(j) = X(j) - \sum_{k=1}^{\max(j-1,q)} L_{X,j,j-k}\, e(j-k) \text{ .}$$

Since $V_{Z,T}$ and $L_{X,T}$ are nested for increasing T then so are $\underset{\sim}{X}_T$ and $\underset{\sim}{e}_T$, i.e. $\underset{\sim}{X}^T_{T+1} = (\underset{\sim}{X}^T_T, X(T+1))$, $\underset{\sim}{e}^T_{T+1} = \left(\underset{\sim}{e}^T_T, e(T+1)\right)$.

With these quantities defined, the algorithm is contained in the following theorem:

Theorem 2:

a) $$Y(T+h|T) = X(T+h|T) - \sum_{j=1}^{p} \alpha(j)\, Y(T+h-j|T)$$

where

i) $$X(T+h|T) = \begin{cases} \sum_{k=h}^{q} L_{X,T+h,T+h-k}\, e(T+h-k) \text{ ,} & h = 1, \ldots, q \\ 0 & h > q \end{cases}$$

ii) $Y(T+h-j|T) = Y(T+h-j)$ if $j \geq h$

b) $$\sigma^2_{T,h} = E\,\{Y(T+h) - Y(T+h|T)\}^2$$

$$= \sum_{k=0}^{h-1} (V^{-1}_{Z,T+h}\, L_{X,T+h})^2_{T+h,T+h-k}\, d_{T+h-k}$$

c) i) $L_{X,T,T-k} \longrightarrow \beta(k)$, $k = 1, \ldots, q$ as $T \longrightarrow \infty$

ii) $(V^{-1}_{Z,T}\, L_{X,T})_{T,T-k} \longrightarrow \beta_\infty(k)$ as $T \longrightarrow \infty$

iii) Let $\gamma(0) = 1, \gamma(1), \gamma(2), \ldots$ be the coefficients of the infinite order moving average representation of the autoregressive part of Y. Then

$$V^{-1}_{Z,p+j,p+j-k} = \gamma(k) , \quad k = 0, 1, \ldots, j, \quad j \geq 0 .$$

iv) $$V^{-1}_{Z,p+j,k} = -\sum_{\ell=1}^{p} \alpha(\ell) V^{-1}_{p+j-\ell,k} , \quad k = 1, \ldots, p-1$$

<u>Proof</u>

Since $\Gamma_{Y,T} \equiv \text{TOEPL}(R_Y(0), \ldots, R_Y(T-1)) = V^{-1}_{Z,T} \Gamma_{X,T} V^{-1}_{Z,T}$ $= V^{-1}_{Z,T} L_{X,T} D_{X,T} L^{T}_{X,T} V^{-T}_{Z,T}$ and the modified Cholesky decomposition is unique, then (b) and (c,ii) follow immediately from Theorem 1. Also, Theorem 1 shows that the elements of the rows of $V^{-1}_{Z,T}$ are converging to the infinite moving average representation of the autoregressive part of Y. Thus the rows of L_X are converging to the infinite moving average representation of the moving average part, that is to $\beta(1), \ldots, \beta(q)$.

To prove (a), note that since $\underset{\sim}{X}_T = V_{Z,T} \underset{\sim}{Y}_T$, we have $\underset{\sim}{\rho}_{XY,T,h}$ $\equiv E(\underset{\sim}{X}_T Y(T+h)) = V_{Z,T} \underset{\sim}{\rho}_{Y,T,h}$. Also for $T+h > p$, $\underset{\sim}{\rho}_{XY,T,h}$ $= E(\underset{\sim}{X}_T [X(T+h) - \sum_{j=1}^{p} \alpha(j) Y(T+h-j)]) = \underset{\sim}{\rho}_{X,T,h} - \sum_{j=1}^{p} \alpha(j) \underset{\sim}{\rho}_{XY,T,h-j}$, $= \underset{\sim}{\rho}_{X,T,h} - \sum_{j=1}^{p} \alpha(j) V_{Z,T} \underset{\sim}{\rho}_{Y,T,h-j}$, where $\underset{\sim}{\rho}_{X,T,h} = E(\underset{\sim}{X}_T X(T+h))$. Thus

$$\underset{\sim}{\rho}_{Y,T,h} = V^{-1}_{Z,T} \underset{\sim}{\rho}_{XY,T,h} = V^{-1}_{Z,T} \underset{\sim}{\rho}_{X,T,h} - \sum_{j=1}^{p} \alpha(j) \underset{\sim}{\rho}_{Y,T,h-j} .$$

Therefore, $Y(T+h|T) = \underset{\sim}{\rho}^{T}_{Y,T,h} V^{T}_{Z,T} L^{-T}_{X,T} D^{-1}_{X,T} \underset{\sim}{e}_T = \underset{\sim}{\rho}^{T}_{X,T,h} V^{-T}_{Z,T} V^{T}_{X,T} L^{-T}_{X,T} D^{-1}_{X,T} \underset{\sim}{e}_T$ $- \sum_{j=1}^{p} \alpha(j) \underset{\sim}{\rho}^{T}_{Y,T,h-j} V^{T}_{Z,T} L^{-T}_{X,T} D^{-1}_{X,T} \underset{\sim}{e}_T = \underset{\sim}{\rho}^{T}_{X,T,h} L^{-T}_{X,T} D^{-1}_{X,T} \underset{\sim}{e}_T$ $- \sum_{j=1}^{p} \alpha(j) Y(T+h-j|T)$.

An argument identical to that used in the proof of Theorem (1a) proves part (i) of (2a). To verify (ii) we substitute for $\underset{\sim}{e}_T$ and $\underset{\sim}{X}_T$ to obtain $\underset{\sim}{\rho}^T_{Y,T,h-j}\, V^T_{Z,T}\, L^{-T}_{X,T}\, D^{-1}_{X,T}\, \underset{\sim}{e}_T = \underset{\sim}{\rho}^T_{Y,T,h-j}\, \Gamma^{-1}_{Y,T} \underset{\sim}{Y}_T = Y(T+h-j)$, since $\underset{\sim}{\rho}^T_{Y,T,h-j}$ is the $(T+h-j)$th row of $\Gamma_{Y,T}$.

To prove (2ciii) we note that comparing coefficients of like powers of z in the equation $1/\sum_{j=0}^{p} \alpha(j)z^j = \sum_{k=0}^{\infty} \gamma(k)z^k$ gives the following recursion for γ: $\sum_{\ell=0}^{k} \alpha(\ell)\gamma(k-\ell) = \delta_k$, $k \geq 0$. Thus we need only show $\sum_{\ell=0}^{k} \alpha(\ell)V^{-1}_{Z,p+j,p+j-(k-\ell)} = \delta_k$, $k = 0, \ldots, j$. But the left hand side of this equation is just the $(p+j)$th row of $V^{-1}_{Z,T}$ times the $(p+j-k)$th column of $V_{Z,T}$. Finally, (2civ) follows by multiplying the $(p+j)$th row of V_Z times the kth column of V_Z^{-1} for $k = 1, \ldots, p-1$.

From Theorem 2 we see that to find $Y(T+h|\Gamma)$ and $\sigma^2_{T,h}$ for $h = h_1, \ldots, h_2$ and $T = T_1, \ldots, T_2$, one essentially needs to calculate V^{-1}_{Z,T_2+h_2} , L_{X,T_2+h_2} , and D_{X,T_2+h_2} . Theorem (2ci) shows that there are q nonzero, nonone elements in a row of L_X and that these elements are converging to the coefficients of the moving average part of Y. Theorem (2ciii) shows that only the first p-1 elements of rows of V^{-1}_{Z,T_2+h_2} are not one of $\gamma(0)$, $\gamma(1)$, ... while (2civ) shows these elements are easily calculated recursively. Thus the number of elements in V^{-1}_{Z,T_2+h_2} and L_{X,T_2+h_2} that need to be calculated and stored in a computer program increases linearly with the number of rows needed prior to attaining convergence. This convergence is illustrated in the next section.

4. A Numerical Example

Consider the autoregressive-moving average process Y of order $p = 4$ and $q = 3$ with $\alpha(1) = -.3357$, $\alpha(2) = .0821$, $\alpha(3) = .1570$, $\alpha(4) = .2567$, $\beta(1) = -.6077$, $\beta(2) = .0831$, $\beta(3) = .1903$, and $\sigma^2 = 1$. Then the variances and first 10 autocorrelations of Y, (denoted $\rho_Y(\cdot)$), the autoregressive part of Y (denoted $\rho_Z(\cdot)$), and the moving average part of Y (denoted $\rho_W(\cdot)$ are given in Table 1, while Table 2 gives the first 10 terms in the infinite order moving average representation of Y, Z, and W.

Table 1. Variances and First 10 Autocorrelations $\rho_Y(\cdot)$, $\rho_Z(\cdot)$, $\rho_W(\cdot)$ of Y, autoregressive part of Y, and moving average part of Y where Y is the above ARMA (4,3) process.

v	$\rho_Y(v)$	$\rho_Z(v)$	$\rho_W(v)$
1	-.2227	.3806	-.4548
2	-.0749	-.0112	-.0230
3	.0616	-.2897	.1347
4	-.1949	-.4128	0
5	-.0015	-.2107	0
6	.0250	.0115	0
7	.0233	.1603	0
8	.0560	.1919	0
9	.0134	.1036	0
10	-.0102	-.0091	0
Variance	1.1306	1.3891	1.4124

Table 2. First 10 Terms in Infinite Order Moving Average Representation of Each of Series in Table 1.

j	ARMA	AR part	MA part
1	-.272	.336	-.608
2	-.090	.031	.083
3	.025	-.174	.190
4	-.198	-.370	0
5	.015	-.201	0
6	.041	-.018	0
7	.037	.114	0
8	.058	.166	0
9	.006	.101	0
10	-.019	.007	0

Finally Table 3 illustrates the convergence proved in Theorem 2 in the above section.

Table 3. The Matrices $V_{Z,10}^{-1} L_{X,10}$, $V_{Z,10}^{-1}$, $L_{X,10}$.

$V_{Z,10}^{-1} L_{X,10}$:

1									
-.223	1								
-.075	-.252	1							
.062	-.064	-.250	1						
-.195	.019	-.091	-.247	1					
-.002	-.205	.013	-.090	-.266	1				
.025	.004	-.206	.013	-.094	-.271	1			
.023	.032	.008	-.206	.021	-.091	-.270	1		
.056	.038	.041	.007	-.200	.024	-.090	-.271	1	
.013	.062	.042	.041	.014	-.199	.025	-.090	-.272	1

$V^{-1}_{Z,10}$:

1									
.381	1								
-.011	.450	1							
-.290	.116	.403	1						
-.413	-.155	.053	.336	1					
-.211	-.389	-.172	.031	.336	1				
.012	-.252	-.382	-.174	.031	.336	1			
.160	-.058	-.226	-.370	-.174	.031	.336	1		
.192	.102	-.031	-.201	-.370	-.174	.031	.336	1	
.104	.178	.112	-.018	-.201	-.370	-.174	.031	.336	1

$L_{X,10}$

1									
-.603	1								
.208	-.702	1							
.338	.102	-.653	1						
	.177	.075	-.583	1					
		.180	.075	-.602	1				
			.180	.077	-.607	1			
				.188	.082	-.606	1		
					.189	.083	-.606	1	
						.190	.083	-.607	1

5. Acknowledgement

The authors would like to thank Professor Emanuel Parzen for helpful discussions. The research was supported in part by Office of Naval Research grant N00014-78-C-0599 and grant CA-28066 from the National Cancer Institute, the HEW.

References

Ali, M.M. (1977). Analysis of autoregressive-moving average models: Estimation and prediction. Biometrika, 64, 535-545.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.

Durbin, J. (1960). The fitting of time-series models. Rev. Int. Statist. Inst., 28, 233-244.

Pagano. M. (1976). On the linear convergence of a covariance factorization algorithm. Journal Ass. Comp. Mach., 23, 310-316.

Whittle, P. (1963). Prediction and Regulation by Linear Least Square Methods. London: The English Universities Press.

Wilkinson, J.H. (1965). The Algebraic Eigenvalue Problem. Oxford: Clarendon Press.

DATE

ILME